https://dx.doi.org/10.22161/ijaers.4.2.34 *ISSN: 2349-6495(P) | 2456-1908(O)*

Privacy Preserving Mining on Vertically Partitioned Database

Anupama K[1], Grace Priya Shalini S[2], Mani.A[3]

[1,2]Department of Computer Science and Engineering, Anna University, Chennai, India
[3]Associative Professor, Department of Computer Science and Engineering, Anna University, India

Abstract— *This system allows multiple data owners to outsource their data in a common could. This paper mainly emphases on privacy preserving mining on vertically partitioned database. It provides an even solution to protect data owner's raw data from the other data owners. To achieve secure outsourcing technique, the system proposes a cloud-aided frequent itemset mining solution. The run time in this system is one order higher than non-privacy preserving mining algorithms.*

Keywords—Association rules, Data Mining, Homomorphic encryption technique, privacy preserving.

I. INTRODUCTION

The major problem in outsourcing data is that it leaks more information about the data owner's data. So outsourcing of data by multiple data owners is not secure. In such a situation, data owners wish to acquire the association rule from a combined dataset and relate a little information about the raw data to other data owners.

Association rule is one of the important analysis techniques that play a vital role in this paper. This paper is built on the Privacy Preserving mining on vertically partitioned databases. This allows the data owners to outsource their data in secured manner. An effective homomorphic encryption technique and could aided mining solution is used.

II. DATA MINING

Data mining is one of the significant study areas in the field of Computer Science. The key objective of the data mining is to extract valuable and significant knowledge from large set of data. Association rule mining is one of the data mining algorithms. Where it provides a solution to determine rules from data. It is also an efficient way to discover similarity among variables in a large database. The study gives different ideas about the several encryption techniques, which can secure data from third parties.

III. HOMOMORPHIC ENCRYPTION

Homomorphic encryption technique is one of the encryption algorithms which is used to perform its operations on cipher texts, when decrypted the result provides the plain text. Additive homomorphic encryption and multiplicative homomorphic are its two types. The addition and multiplication is to be performed on the cipher texts.

IV. SYMMETRIC HOMOMORPHIC ENCRYPTION

The symmetric homomorphic encryption technique uses the private key to encrypt the data and it uses many homomorphic additions and limited number of multiplications.

V. KEY GENERATION ALGORITHM

Key generation algorithm is the process of generating keys in cryptography. The generated key is used for encrypting and decrypting the data.

VI. ENCRYPTION ALGORITHM

It is a mathematical procedure for performing encryption on data while encrypting, we get a meaningless cipher text. And a key is used to transfer data to original form.

VII. DECRYPTION ALGORITHM

Decryption algorithm involves the process of converting encoded or encrypted text to its original form. This algorithm requires a secret key for the conversion.

VIII. SYSTEM MODEL

The system model mainly comprises of two or more data owners and a cloud. Each data owners have its own private database and data owners are used to encrypt the database to the cloud. The main task of the cloud is storing and compiling the received database.

https://dx.doi.org/10.22161/ijaers.4.2.34 *ISSN: 2349-6495(P) | 2456-1908(O)*

Fig. 1: System model of outsourced data

IX. LITERATURE SURVEY

Boxiang Dong [1], provides an capable mechanisms to verify the result reliability of outsourced data mining computations. They emphases on importance of data mining problem and data mining service. Their idea is to construct frequent item sets from real items, so first they remove real items from the original data set to construct artificial item sets. One of the nice property is frequent item set are independent from the size of the data sets and also they are used to untrusted server that may try to escape. They are used to remove small set of items from large data sets.

Junzuolai [2], concentrates on data mining study with enormous application such as network intrusion detection technique. Here data owner need to encrypt the original data using suitable algorithm, it can be performed by cloud serves on the encrypted data directly.

FoscaGiannotti [3], proposed an attack model based on background knowledge and provide a scheme for privacy preserving mining. They adopt a Frequency based attack model in which the server knows the exact set of items. These attack models helps to solve the problem of outsourcing the association rule mining task.

Md. GolamKaosar[4], elaborates the homomorphic encryption technique that produces accurate result. Using only fully-homomorphic encryption they represented two-party association rule mining algorithm. This algorithm mainly used to calculate the support and confidence of association rule and it will return a single bit.

P.Fournier-Viger[5], the main is to discover interesting patterns and association in database. Mostly open source data mining such as Weka (Witten et al) provides a wide range of mining technique. It also offers limited set of algorithms. Each algorithm provides a sample input file.

BamshadMobasher[6], the main is to improve the scalability of patterns, while improving the overall effectiveness . Here they mainly use KNN approach (K-nearest neighbour approach). By this KNN approach effectiveness of the system is measured. They have presented a scalable framework formed on association rule mining. The framework includes a resourceful data structure for storing frequent item sets.

W.K.Wong[7], the proposed substitution cipher technique to get a relief from high mining cost, minimization of demand, mining for multiple Owners and for security. Substitution technique is one the highly secure with a low data transformation cost. Transformation cost is one the method for saving global mining in distributed manner.

Tom Brijs[8], elaborated the method of KDD (Knowledge Discovery Database). This model enables the integration of qualitative standards. They use frequent item set to demonstrate profitability in easy and sensitive way. They are used to optimise the problem and therefore reflect framework of decision maker.

Selim V. Kaya[9], planned a privacy preservative distributed data mining techniques are developed for creating a data mining model over distributed database. They are mainly used for make high effectiveness in communication and computation in distributed data mining.

N.V Muthu Lakshmi [10], introduced a cryptography technique to discover hidden information from large database. Here more attention is towards association rule mining between item and set of items. They provide a distributed database environment and also play an important role in problem definition.

X. PROPOSED SYSTEM

The system proposes an efficient symmetric homomorphic encryption technique scheme.By this homomorphic encryption scheme; a secure outsourcing process is achieved. This allows multiple data owner to outsource their data to the other owners. The symmetric homomorphic encryption technique uses many homomorphic additions and limited number of multiplications.

ADVANTAGES

- High effectiveness and scalability.
- Allows multiple data owners to outsource in common form.
- Run time of each solution is one order greater.
- Leak less information.
- Low complexity.

XI. CONCLUSION

In this paper we propose mining solution for vertically partitioned database .In this multiple data owners are allowed to outsource their data. The obtained solution

https://dx.doi.org/10.22161/ijaers.4.2.34 *ISSN: 2349-6495(P) | 2456-1908(O)*

protects the data owner's data from other third parties. This also ensures securing the mining result. This solution outflows a lesser amount of information about the data owner's data when compared with other solutionand it also very efficient.

REFERENCES

[1] Boxiang Dong, Ruilin Liu, Hui(Wendy) Wang "Result Integrity Verification of Outsourced Frequent Itemset Mining" in *Proc. SIAM Int. Conf. Data Mining*, Vancouver, BC, Canada, Apr./May 2015

[2] J. Lai, Y. Li, R. H. Deng, J. Weng, C. Guan, and Q. Yan, "Towards semantically secure outsourcing of association rule mining on categorical data," inf. sci., vol.267, pp 267-286 May 2014.

[3] FoscaGiannotti, Laks V. S. Lakshmanan, Anna Monreal, Dino Pedreschi, and Hui(Wendy) Wang "Privacy Preserving Mining of Association Rule From Outsourced Transaction database" vol 7, no.3, September 2013.

[4] Md. GolamKaosar, Russell PauletXun Yi, "Secure Two-Party Association Rule Mining" in Proc ACSW-AISC, pp. 15-22, 2011.

[5] P.Fournier-Viger, A.Gomariz, T.Gueniche, A.Soltani, C.W.Wu, V.S.Tseng, "SPMF: java Open-Source pattern mining library", vol.15, pp.3389-3393, 2014.

[6] Bamshad Mobasher, Honghua Dai, Tao Luo, Miki Nakagawa "Effective Personalised Based on Association Rule Discovery from Web Usage Data" ,Proc WIDM, pp 9-15, 2001

[7] W.K.Wong, David W.Cheng, Edward Hung, Ben Kao, Nikos Mamoulis,"Security in Outsourcing of Association Rule Mining" 2007

[8] Tom Brijs, Gilbert Swinnen, KoenVanHoof, Geert Wets, "Using Association Rule for Product Assortment Decisions: A case study" SIGKDD, 1999, pp 254-260.

[9] Selim V. Kaya, Thomas B. Pedersen, ErkaySavas and YucelSaygun, "Efficient privacy preserving Distributed Clustering Based on Secret Sharing", pp 280-291, 2007

[10] N.V Muthu Lakshmi, Dr. K Sandhya Rani " Privacy Preserving Association Rule Miing in Horizontally Partitioned Database using Cryptography technique" vol.3, pp 3176-3183, 2012